

**Implementasi Data Mining Untuk Menentukan Kelayakan Penerimaan Bedah Rumah Menggunakan Algoritma *K-Means* Fakultas Ilmu Komputer Universitas Pasir Pengaraian**

**Dona<sup>1</sup>, Khairul Sabri<sup>2</sup>, Kiki Yasdomi<sup>3</sup>, Mi'rajul Rifqi<sup>4</sup>, Chandra Yuliansyah<sup>5</sup>, Afnes Susmita<sup>6</sup>**

Ilmu Komputer, Universitas Pasir Pengaraian<sup>1,2,3,4,6</sup>  
Fakultas Teknik, Universitas Pat Petulai<sup>5</sup>

Email: [dona201804@gmail.com](mailto:dona201804@gmail.com)<sup>1</sup>, [khairulsabri.tch@gmail.com](mailto:khairulsabri.tch@gmail.com)<sup>2</sup>, [kikiyasdomi@gmail.com](mailto:kikiyasdomi@gmail.com)<sup>3</sup>, [mirajulrifqi@gmail.com](mailto:mirajulrifqi@gmail.com)<sup>4</sup>, [chandrayuliansyah1987@gmail.com](mailto:chandrayuliansyah1987@gmail.com)<sup>5</sup>, [afnessusmita20@gmail.com](mailto:afnessusmita20@gmail.com)<sup>6</sup>

*Corresponding Author* : [dona201804@gmail.com](mailto:dona201804@gmail.com)

***Abstract***

*The house renovation program is a government program, namely the Public Housing and Settlement Area Office, which is shown to rehabilitate or repair community houses that are not habitabel. This program is often a problem in Bangun Purba Village because when determining house renovation assistance, it is still by looking or relying on thoughts and predictions for people who will get assistance, so that the assistance that will be distributed is not on target. This house renovation program aims to provide welfare for the poor to build or renovate houses that are not habitabel. Therefore, the author created a Data Mining system to determine the feasibility of receiving home renovation using the K-Means algorithm using the PHP programming language and MySQL database. The K-Means method utilizes training data to generate probabilities such as each grouping for different classes, so that the probability values of the grouping can be optimized to determine the eligibility for receiving home surgery assistance based on the grouping process carried out by the K-Means method.*

***Keywords*** : *Data Mining, Eligibility for House Surgery, K-Means Algorithm.*

**I. Pendahuluan**

Perumahan merupakan kebutuhan dasar manusia yang harus dipenuhi dalam rangka mewujudkan kehidupan yang layak dan sejahtera. Menurut Undang-Undang Nomor 1 Tahun 2011 tentang Perumahan dan Kawasan Permukiman, Pasal 1 Ayat 7, "Rumah adalah bangunan gedung yang

berfungsi sebagai tempat tinggal yang layak huni, sarana pembinaan keluarga, cerminan harkat dan martabat penghuninya, serta aset bagi pemilikinya".

Salah satu upaya pemerintah dalam memenuhi kebutuhan tersebut adalah melalui program bedah rumah, yang merupakan program rehabilitasi rumah tidak layak huni

(RTLH) untuk masyarakat miskin. Program ini didanai oleh Dinas Perumahan Rakyat dan Kawasan Permukiman, dengan memberikan bantuan bangunan serta menyediakan tenaga tukang dan teknisi untuk merenovasi rumah yang tidak layak huni (Chamida et al., 2021).

Namun demikian, pelaksanaan program ini seringkali menghadapi permasalahan dalam proses penentuan penerima bantuan. Di Desa Bangun Purba, penetapan penerima masih dilakukan secara subjektif, yaitu dengan mengandalkan pengamatan dan penilaian manual oleh pihak desa. Hal ini menyebabkan bantuan tidak tersalurkan secara merata dan tepat sasaran, serta memunculkan potensi ketidakadilan dalam penentuan penerima bantuan.

Dalam proses penyaluran bantuan tersebut, diperlukan suatu sistem pendukung keputusan yang mampu menilai kelayakan secara objektif berdasarkan data dan kriteria yang jelas. Kriteria umum yang digunakan mencakup pendapatan, jumlah tanggungan, pekerjaan, jenis atap, jenis lantai, jenis dinding, dan aspek kepemilikan rumah. Beberapa penelitian sebelumnya juga menambahkan kriteria seperti kondisi toilet, luas tanah kurang dari 8 m<sup>2</sup>, dan kondisi dinding dalam ruangan sebagai indikator rumah tidak layak huni.

Menurut Kusnadi & Putri (2021), seiring dengan perkembangan teknologi informasi, data mining menjadi solusi yang menjanjikan untuk menangani permasalahan ini. Data mining merupakan proses penggunaan teknik statistik, kecerdasan buatan, dan machine learning untuk mengekstraksi informasi penting dari kumpulan data yang besar. Salah satu algoritma dalam data mining yang dapat

diterapkan untuk proses pengelompokan data adalah K-Means *Clustering*.

K-Means adalah algoritma yang membagi data ke dalam sejumlah kelompok (cluster) berdasarkan tingkat kemiripan, dengan menggunakan jarak Euclidean sebagai pengukuran utama. Algoritma ini menentukan pusat cluster awal, lalu mengelompokkan data berdasarkan kedekatannya terhadap pusat tersebut. Proses ini diulang hingga hasil pengelompokan stabil (Pangestu & Ridwan, 2022).

Penelitian ini bertujuan untuk membangun sistem pendukung keputusan berbasis data mining dengan algoritma K-Means guna menentukan kelayakan penerima bantuan bedah rumah secara objektif dan efisien di Desa Bangun Purba. Dengan adanya sistem ini, diharapkan program bantuan dapat tersalurkan secara adil, tepat sasaran, dan transparan.

## II. Landasan Teori

Jika Teori anda melebihi satu, gunakan level kedua teoriseperti di bawah ini. Pengutipan dilakukan dengan mencantumkan. Nama penulis dan tahun penerbitan mengikuti kalimat. Penomoran harus diberikan pada gambar (gambar, grafik, foto dan peta), tabel dan persamaan matematika, rekasi kimia dan fisika berdasarkan urutannya.

### ***Knowledge Discovery in Database (KDD)***

*Knowledge Discovery in Database (KDD)* merupakan metode untuk memperoleh pengetahuan dari database yang ada. Di dalam database terdapat tabel – tabel yang saling berhubungan/berelasi. Hasil dari proses KDD dapat digunakan

sebagai data dasar untuk pemrosesan lebih lanjut (Srirahayu & Pribadie, 2023).

Proses *Knowledge Discovery in Database (KDD)* termasuk istilah data mining, yang mengubah data mentah menjadi lebih bermanfaat dan berguna (Damanik et al., 2021).

1. Data Cleansing

Data Cleansing adalah fase di mana data yang tidak lengkap, tidak konsisten, dan tidak tepat dibuang dari koleksi data. Ini memungkinkan data yang relevan digunakan untuk penggalian pengetahuan (discovery knowledge).

2. Data Integration

Pada tahap ini sumber data yang berulang (multiple data) dan file yang berulang (multiple file) dapat digabungkan dan digabungkan ke dalam satu sumber.

3. Selection

Pada tahap ini, data yang relevan untuk analisis dapat dipilih dan diambil dari koleksi data yang ada.

4. Data Transformation

Tahap ini disebut konsolidasi data. Pada tahap ini, data yang terpilih akan dinormalisasi dan digabungkan sehingga sesuai dengan prosedur penggalian.

5. Data Mining

Tahap ini adalah yang paling penting karena proses yang digunakan untuk mengekstrak pola-pola yang mungkin berguna.

6. Pattern Evaluation

Pada titik ini, pola-pola yang menarik dengan jelas menunjukkan

informasi yang telah ditemukan berdasarkan ukuran yang telah diberikan.

7. *Knowledge Representation*

Tahap terakhir menunjukkan hasil data mining secara visual kepada pengguna. Ini menggunakan visualisasi untuk membantu pengguna memahami dan memvisualisasikan hasilnya (Damanik et al., 2021).

### **Data Mining**

Data Mining adalah suatu proses pencarian data secara otomatis dapat mendapatkan sebuah model dari database yang besar. Data Mining merupakan metode untuk menemukan informasi tersembunyi dalam database dan bagian dari proses *Knowledge Discovery in Databases (KDD)* untuk menemukan informasi dan pola yang berguna dalam data (Dewi et al., 2022).

### **Data Preprocessing**

Data Preprocessing merupakan kegiatan pemrosesan data untuk mengubah atau menyandikan data sehingga dapat dengan mudah diinterpretasikan oleh mesin. Tahapannya adalah cleaning data, transformasi data dan normalisasi data (Rahayu et al., 2023).

### **Clustering**

*Clustering* adalah teknik pengelompokan objek berdasarkan kesamaan antar objek. *Clustering* dapat mengelompokkan data menjadi satu cluster dengan mempertimbangkan nilai kemiripan maksimum dan minimum. Nilai pada metode

ini difokuskan untuk data bertipe numerik. Pada *Clustering* terdapat istilah anomali. Anomali merupakan data yang telah di uji pada *Clustering* dan tidak menemukan kelas yang cocok (Pangestu & Ridwan, 2022).

### Bedah Rumah

Bedah rumah adalah kegiatan yang dilakukan untuk memperbaiki rumah tidak layak huni, yaitu rumah yang tidak memenuhi syarat kesehatan, keamanan, dan sosial. Bantuan bedah rumah diberikan sebagai cara untuk membantu orang yang tidak mampu memperbaiki rumah atau tempat tinggal mereka sendiri (Pasaribu, 2022).

### K-Means

*K-Means* merupakan teknik *Clustering* yang diperoleh dari sebuah dataset dengan cara menghitung jarak dari setiap titik ke pusat *cluster* secara *iterative* (Prastiwi et al., 2022).

Langkah-langkah melakukan *Clustering* dengan metode *K-Means* adalah sebagai berikut (Silalahi, 2020):

1. Pilih jumlah cluster k.
2. Inisialisasi k pusat cluster ini bisa dilakukan dengan berbagai cara. Namun yang paling sering dilakukan adalah dengan cara random. Pusat-pusat cluster diberi nilai awal dengan angka-angka random.
3. Alokasikan semua data/ objek cluster terdekat. Kedekatan kedua objek ditentukan berdasarkan jarak kedua objek tersebut. Demikian juga kedekatan suatu data ke cluster tertentu ditentukan jarak antara data dengan pusat cluster.

Dalam thap ini perlu dihitung jarak tiap data ke tiap pusat cluster. Jarak paling antara satu data dengan satu cluster tertentu akan menentukan suatu data masuk dalam cluster yang mana. Untuk menghitung jarak semua data ke setiap titik pusat cluster dapat menggunakan teori jarak Euclidean yang dirumuskan sebagai berikut :

$$D(i,j) = \frac{=}{\sqrt{(x_{1i} - x_{1j})^2 + (x_{2i} - x_{2j})^2 + \dots + (x_{ki} - x_{kj})^2}}$$

Dimana

$D(i,j)$  = Jarak data ke i ke pusat kluster j

$X_{ki}$  = Data ke i pada atribut data ke k

$X_{kj}$  = Titik pusat ke j pada atribut ke k

- a. Hitung kembali pusat cluster dengan keanggotaan cluster yang sekarang. Pusat cluster adalah rata-rata dari semua data/ objek dalam cluster tertentu. Jika dikehendaki bisa juga menggunakan median dari cluster tersebut. Jika rata-rata (mean) bukan satu-satunya ukuran yang bisa di pakai.
- b. Tugaskan lagi setiap objek memakai cluster yang baru. Jika pusat cluster tidak berubah lagi maka proses *Clustering* selesai. Atau, kembali ke langkah nomor 3 sampai pusat cluster tidak berubah lagi.

### **Bahasa Pemrograman**

Bahasa pemrograman, juga dikenal sebagai "bahasa komputer" atau "bahasa pemrograman komputer", adalah protokol umum untuk mengelola komputer. Ini terdiri dari set aturan sintaks dan semantik yang digunakan untuk membuat program komputer (Premana et al., 2022).

### **HTML**

HTML (*Hyper Text Markup Language*) adalah skrip yang terdiri dari tag-tag yang digunakan untuk membuat dan mengatur struktur website. HTML (*Hyper Text Markup Language*) adalah bahasa pemrograman yang digunakan untuk membuat halaman web (Suhartini et al., 2020).

### **Cascading Style Sheets (CSS)**

CSS adalah singkatan dari Cascading Style Sheet. Kegunaan CSS yaitu untuk mengatur tampilan dokumen HTML, sebagai contoh pengaturan jarak antar baris, teks, format border, warna bahkan hingga penampilan file gambar (Rudjiono & Saputro, 2021).

### **Java Script**

Java Script adalah script program berbasis client yang di eksekusi oleh browser sehingga membuat halaman web melakukan tugas-tugas tambahan yang tidak bisa dilakukan oleh script HTML biasa (Noviantoro et al., 2022).

### **Structured Query Language (SQL)**

SQL adalah sebuah konsep pengoperasian basis data terutama untuk

proses seleksi, pemasukan, pengubahan, dan penghapusan data yang dimungkinkan dapat dikerjakan dengan mudah dan otomatis (Pangestu & Ridwan, 2022).

### **Website**

Website adalah salah satu media yang sangat berguna bagi kehidupan manusia sekarang ini, baik untuk sarana promosi, pembelajaran, komunikasi dan masih banyak lagi sarana yang menggunakan website (Rudjiono & Saputro, 2021).

### **III. Metode Penelitian**

Kerangka kerja yang akan dilakukan pada penelitian ini dapat diuraikan seperti berikut ini:

#### **1. Identifikasi Masalah**

Pada tahap ini diidentifikasi masalah yang akan menjadi aspek penelitian. Identifikasi masalah didapat melalui pengamatan secara langsung terhadap objek ini dan dilakukan dengan maksud agar dapat mengetahui secara jelas permasalahan yang terkait dengan sistem yang akan dirancang. Setelah diamati permasalahan yang menjadi masalah utama yaitu belum adanya sistem *Clustering* kelayakan penerima badah rumah yang dibutuhkan untuk menentukan kelayakan penerima bantuan.

#### **2. Analisa Masalah**

Analisa masalah adalah langkah untuk memahami masalah yang telah ditentukan di ruang lingkup dan batasannya. Dengan menganalisa masalah yang telah ditentukan tersebut, maka perlu dibuat sebuah sistem. Sistem ini bertujuan untuk membantu aparat desa di Desa Bangun

Purba untuk menentukan kelayakan penerima bedah rumah.

### 3. Study Literatur

Setelah menganalisa masalah, maka dipelajari literatur yang berhubungan dengan permasalahan. Kemudian literatur yang dipelajari dan diseleksi supaya dapat menentukan literatur yang akan digunakan dalam penelitian. Kemudian literatur yang berhubungan dengan permasalahan, kemudian literatur yang dipelajari tersebut diseleksi untuk ditentukan literatur yang akan digunakan pada penelitian ini. Literatur didapatkan dari jurnal yang membahas tentang Kelayakan Penerimaan Bedah Rumah.

### 4. Pengumpulan Data

Pengumpulan data dilakukan untuk memperoleh informasi yang dibutuhkan dalam penelitian. Data yang dibutuhkan berupa:

#### a. Data Primer

Dari hasil wawancara dengan Kepala Desa Bangun purba beserta Staf diperoleh hasil data berupa data masyarakat yang mengurus bantuan bedah rumah. Selain itu, diketahui bahwa penentuan calon penerima bantuan bedah rumah masih dilakukan dengan manual serta belum adanya sistem sehingga proses penentuan calon penerima bantuan bedah rumah tidak efektif dan tidak tepat sasaran.

#### b. Data Sekunder

Pada penelitian ini, data sekunder yang diperoleh merupakan jurnal-jurnal dan data pendukung yang digunakan sebagai referensi pada penelitian yang dilakukan.

### 5. Penerapan Algoritma K-Means

Pada tahap ini adalah melakukan perhitungan secara manual dengan

Algoritma K-Means yang berkaitan dengan kasus yang diteliti dan apa-apa saja data yang dibutuhkan oleh peneliti untuk mencegah permasalahannya.

### 6. Perancangan Sistem

Perancangan sistem meliputi bagaimana kegiatan dalam siklus pengembangan sistem bisa diterapkan secara efektif dan efisien sehingga mampu menghasilkan sebuah sistem yang sesuai dengan tujuan. Perancangan sistem ini dibuat dalam bentuk Analisa Sistem Informasi (ASI), *Context Diagram*, *Data Flow Diagram* (DFD), *Entity Relationship Diagram* (ERD), *Flowchart*.

### 7. Pembuatan Program

Pada tahap ini adalah tahap design sistem akan diterjemahkan kedalam bahasa pemrograman. Yang mana bahasa pemrograman yang digunakan dalam pembuatan sistem ini adalah bahasa pemrograman PHP dengan database MySQL.

### 8. Pengujian Sistem

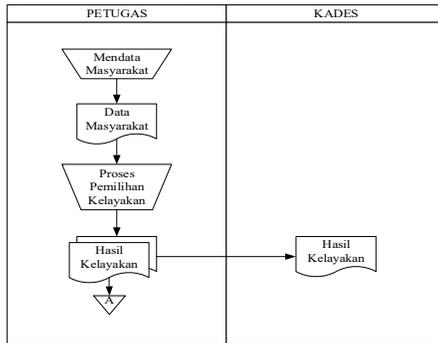
Pada tahap ini adalah tahap pengujian yang dilakukan dengan menggunakan metode Black Box, hal ini dilakukan untuk meminimalisir terjadinya kesalahan dan yang dihasilkan sesuai dengan yang diinginkan.

## IV. Hasil dan Pembahasan

### Analisa Sistem

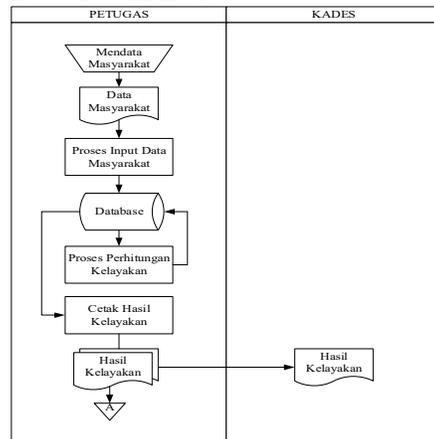
Analisa sistem dilakukan untuk memecahkan proses *Clustering* penerima bantuan bedah rumah menggunakan metode *K-Means*, yang mana outputnya menghasilkan suatu keputusan layak dan tidak layak sebagai penerima bantuan bedah rumah.

### Analisa Sistem Lama



Gambar 1. Analisa Sistem Lama

### Analisa Sistem Baru



Gambar 2. Analisa Sistem Baru

### Analisa Data

Tahap ini diperlukan data masyarakat untuk dapat diproses menggunakan metode K-Means yang menggunakan beberapa atribut yang didapat dari data tersebut.

### Data Masyarakat

Berdasarkan sumbernya, data penelitian yang diperoleh adalah data masyarakat Bangun Purba. Data yang diperoleh dari Kantor Desa Bangun Purba.

### Data Selection Data Masyarakat

Data masyarakat yang telah diperoleh selanjutnya adalah melakukan proses seleksi data dimana data yang tidak digunakan dibuang/dihilangkan. Data yang digunakan seperti : No, Desa/Kelurahan, Alamat, Kepala Keluarga, NIK, Jenis Kelamin, Tanggal Lahir, Pekerjaan, Pendidikan, Jumlah Pendapatan, Jumlah Tanggungan, Kepemilikan Rumah, Jenis Atap, Jenis Dinding, Jenis Lantai, Sumber Penerangan, Bahan Bakar Memasak, Sumber Air Minum, Penerima BPNT, Penerima BST, Penerima PKH, Penerima Sembako, Resiko Stanting. Data masyarakat yang digunakan akan membentuk data selection seperti : No, Kepala Keluarga, Pekerjaan, Jumlah Pendapatan, Jumlah Tanggungan, Kepemilikan Rumah, Jenis Dinding, Jenis Lantai.

### Data Cleaning Data Masyarakat

Pada tahap ini, data yang telah diseleksi selanjutnya melakukan penghapusan data, dan memperbaiki kesalahan data. Berikut ini merupakan beberapa isi data yang telah dibersihkan (data cleaning).

### Pra-Proses Data

Sebelum melakukan transformasi data, terlebih dahulu melakukan pra-proses data berdasarkan atribut yang telah di seleksi. Data masyarakat yang ada kemudian digunakan sebagai data yang diolah. Penyesuaian format data dilakukan untuk mempermudah pengolahan data. Format data disesuaikan sehingga menjadi bentuk angka.

### 1. Clustering Pekerjaan

Pembentukan *Clustering* pada *Clustering* Pekerjaan akan dilakukan proses pembobotan. Semakin rendah tingkat pekerjaan semakin rendah pembobotannya, begitu pula sebaliknya semakin tinggi tingkat pekerjaan semakin tinggi pula pembobotannya. Berikut pembobotan pada Pekerjaan:

**Tabel 1. Clustering Pekerjaan**

Pekerjaan	Bobot
Petani	1
Pekebun	2
Buruh Lepas	3
Pedagang	4
Wiraswasta	5

### 2. Clustering Jumlah Pendapatan

Pembentukan *Clustering* pada *Clustering* Jumlah Penghasilan akan di proses menggunakan distribusi frekuensi dengan beberapa langkah sebagai berikut :

#### a. Menentukan Jumlah Kelas (K)

Berdasarkan data masyarakat yang telah di Cleaning maka akan ditemukan jumlah kelas, yaitu dengan rumus STURGESS, yang formulasinya sebagai berikut:

$$\text{Jumlah Kelas (K)} = 1 + 3,3 \log n$$

Dimana n = jumlah data yang dimiliki

Berikut ini perhitungan Jumlah Kelas dengan rumus tersebut sehingga diperoleh:

$$K = 1 + 3,3 \log 106 = 1 + 6,683509355358 = 7,683509355358 \text{ dibulatkan menjadi } 8 \text{ Kelas}$$

Jumlah data yang dimiliki (n), berdasarkan jumlah keseluruhan data masyarakat.

#### b. Menentukan Range

Range dapat diartikan sebagai jarak antara data terkecil sampai ke terbesar atau selisih antara data terbesar sampai terkecil. Dari data masyarakat yang telah didapatkan

maka untuk mendapatkan range dilakukan sebagai berikut :

$$\begin{aligned} \text{Range (R)} &= \frac{\text{Jumlah Pendapatan Terbesar} - \text{Jumlah Pendapatan Terkecil}}{\text{Banyak Kelas}} \\ &= \frac{2.000.000 - 500.000}{8} \\ &= \frac{1.500.000}{8} \\ &= 187,500 \end{aligned}$$

Berdasarkan hasil perhitungan atribut jumlah pendapatan menggunakan distribusi frekuensi maka *Clustering* dapat dilihat pada tabel 4

**Tabel 2. Clustering Jumlah Pendapatan**

Clustering Jumlah Pendapatan	Kelas (K)	Range(R)	Bobot
	1	500.000-687,500	1
Rendah	2	687,500-875.000	
	3	875.000-1.062.500	
	4	1.062.500-1.250.000	
	5	1.250.000-1.437.500	2
Tinggi	6	1.437.500-1.625.000	
	7	1.625.000-1.812.500	
	8	1.812.500-2.000.000	

Berdasarkan tabel 4, kelas 1, kelas 2, kelas 3, kelas 4, dikategorikan sebagai *Clustering* rendah dengan bobot 1, sedangkan kelas 5, kelas 6, kelas 7, kelas 8 dikategorikan sebagai *Clustering* tinggi dengan bobot 2. Perolehan *Clustering* pada Tabel 4.3 ditetapkan oleh pihak terkait setelah dilakukan

perhitungan menggunakan distribusi frekuensi.

**Clustering Jumlah Tanggungan**

Pembentukan *Clustering* pada *Clustering* Jumlah Penghasilan akan di proses menggunakan distribusi frekuensi dengan beberapa langkah sebagai berikut :

a. Menentukan Jumlah Kelas (K)

Berdasarkan data masyarakat yang telah di Cleaning maka akan ditemukan jumlah kelas, yaitu dengan rumus STURGESS, yang formulasinya sebagai berikut:

$$\text{Jumlah Kelas (K)} = 1 + 3,3 \log n$$

Dimana n = jumlah data yang dimiliki

Berikut ini perhitungan Jumlah Kelas dengan rumus tersebut sehingga diperoleh:

$$K = 1 + 3,3 \log 106 = 1 + 6,683509355358 = 7,683509355358$$

dibulatkan menjadi 8 Kelas

Jumlah data yang dimiliki (n), berdasarkan jumlah keseluruhan data masyarakat.

b. Menentukan Range

Range dapat diartikan sebagai jarak antara data terkecil sampai ke terbesar atau selisih antara data terbesar sampai terkecil. Dari data masyarakat yang telah didapatkan maka untuk mendapatkan range dilakukan sebagai berikut :

$$\begin{aligned} \text{Range (R)} &= \text{Jumlah Tanggungan} \\ &\quad \text{Terbesar} - \text{Jumlah} \\ &\quad \text{Tanggungan} \\ &\quad \text{Terkecil} / \text{Banyak Kelas} \\ &= 5 - 2 / 8 \\ &= 3 / 8 \end{aligned}$$

= 0,375 dibulatkan menjadi 0 range

Berdasarkan hasil perhitungan atribut jumlah pendapatan menggunakan distribusi frekuensi maka *Clustering* dapat dilihat pada tabel 5

**Tabel 3. Clustering Jumlah Tanggungan**

<i>Clustering</i> Jumlah Tanggungan	Kelas (K)	Range(R)	Bobot
Sedikit	1	2	2
	2	3	
Banyak	3	4	1
	4	5	
	5	6	
	6	7	
	7	8	
	8	9	

**3. Clustering Kepemilikan Rumah**

Pembentukan *Clustering* pada *Clustering* Kepemilikan Rumah akan dilakukan proses pembobotan. Kepemilikan rumah milik sendiri lebih diutamakan dari pada yang lainnya karena kemilikan rumah dengan milik sendiri merupakan salah satu syarat untuk menerima bantuan bedah rumah. Jadi dengan kepemilikan rumah milik sendiri diberi bobot 1. Berikut pembobotan pada kepemilikan rumah

**Tabel 4. Clustering Kepemilikan Rumah**

Kepemilikan Rumah	Bobot
Milik Sendiri	1
Menumpang	2

#### 4. Clustering Jenis Dinding

Pembentukan *Clustering* pada *Clustering* Jenis Dinding akan dilakukan proses pembobotan. Semakin rendah tingkat jenis dinding semakin rendah pembobotannya, begitu pula sebaliknya semakin tinggi tingkat jenis dinding semakin tinggi pula pembobotannya. Berikut pembobotan pada jenis dinding :

**Tabel 5. Clustering Jenis Dinding**

Jenis Dinding	Bobot
Papan	1
Tembok	2
Semen	3

#### 5. Clustering Jenis Lantai

Pembentukan *Clustering* pada *Clustering* Jenis Lantai akan dilakukan proses pembobotan. Semakin rendah tingkat jenis lantai semakin rendah pembobotannya, begitu pula sebaliknya semakin tinggi tingkat jenis lantai semakin tinggi pula pembobotannya. Berikut pembobotan pada jenis lantai :

**Tabel 6. Clustering Jenis Lantai**

Jenis Lantai	Bobot
Papan	1
Semen	2
Keramik	3

#### Data Transformasi Data Masyarakat

Setelah clustering penilaian ditentukan, data masyarakat harus di transformasi ke dalam bentuk angka, berikut hasil transformasi teks ke dalam angka yang sudah ditentukan.

#### Iterasi ke-1

Pada iterasi pertama dimulai dengan menentukan nilai centroid awal

dengan mengambil nilai secara acak pada tabel transformasi pada perhitungan diatas adalah pada data ke-2 dan data ke-12 setelah diperoleh pusat cluster 1 dan cluster 2, selanjutnya dilakukan perhitungan jarak dari data terhadap pusat cluster 1 dan cluster 2 dengan menggunakan Euclidian Distance, maka diperoleh hasil untuk nilai tertinggi diinisialkan nilai clusternya adalah = 0, sedangkan untuk nilai cluster terendah diinisialkan adalah = 1.

Pada pencarian iterasi pertama didapatkan hasil sebagai berikut :

1. (data ke-1, data ke-2, data ke-3, data ke-5, data ke-7, data ke-8, data ke-9, data ke-11, data ke-13, data ke-15, data ke-16, data ke-18, data ke-20, data ke-23, data ke-24) = 15 anggota C1 (Layak)
2. (data ke-4, data ke-6, data ke-10, data ke-12, data ke-14, data ke-17, data ke-19, data ke-21, data ke-22, data ke-25) = 10 anggota C2 (Tidak Layak)

#### Iterasi ke-2

Menentukan posisi centroid baru dengan cara menghitung nilai rata-rata dari data yang ada pada setiap centroid yang sama.

Pada iterasi kedua didapatkan hasil sebagai berikut :

1. (data ke-1, data ke-2, data ke-3, data ke-5 data ke-8, data ke-9, data ke-13, data ke-15, data ke-20, data ke-23, data ke-24) = 11 anggota C1 (Layak)
2. (data ke-4, data ke-6, data ke-7, data ke-10, data ke-11, data ke-12, data ke-14, data ke-16, data ke-17, data ke-18, data ke-19, data ke-21, data



ID	Nama Penerima	Pekerjaan	Jumlah Pendapatan	Jumlah Tanggungan	Kepemilikan Rumah	Jenis Dinding	Jenis Lantai	
1	Alpaudin	2	1	2	2	1	2	Detail
2	Sudirman	1	1	2	2	2	2	Detail
3	Panuhum Hasbiuan	1	2	2	2	2	2	Detail
4	Sanal Akhlaqy	4	2	2	2	2	2	Detail
5	Burhan	1	1	2	2	2	2	Detail
6	Henni	4	2	2	2	2	2	Detail
7	Fahanta	3	1	1	2	2	2	Detail
8	Arhani	1	1	2	2	2	2	Detail

**Gambar 8. Tampilan Halaman Data Nilai Bobot**

ID	Nama Penerima	Pekerjaan	Jumlah Pendapatan	Jumlah Tanggungan	Kepemilikan Rumah	Jenis Dinding	Jenis Lantai	Cluster
1	Alpaudin	2	1	2	2	1	2	Cluster
2	Sudirman	1	1	2	2	2	2	Cluster
3	Panuhum Hasbiuan	1	2	2	2	2	2	Cluster
4	Sanal Akhlaqy	4	2	2	2	2	2	Cluster
5	Burhan	1	1	2	2	2	2	Cluster
6	Henni	4	2	2	2	2	2	Cluster
7	Fahanta	3	1	1	2	2	2	Cluster
8	Arhani	1	1	2	2	2	2	Cluster

**Gambar 9. Tampilan Halaman Data Penilaian**

ID	Nama Penerima	Pekerjaan	Jumlah Pendapatan	Jumlah Tanggungan	Kepemilikan Rumah	Jenis Dinding	Jenis Lantai	Cluster 1	Cluster 2
1	Alpaudin	2	1	2	2	1	2	L14213523731	3,46410161532
2	Sudirman	1	1	2	2	2	2	0	4,47213395499
3	Panuhum Hasbiuan	1	2	2	2	2	2	1	4,35880843541

**Gambar 10. Tampilan Halaman Detail Proses Pengelompokan**

## V. Kesimpulan dan Saran

**Kesimpulan**  
Berdasarkan uraian dan pembahasan pada bab-bab sebelumnya, maka dapat disimpulkan bahwa:

1. Pada penelitian ini melakukan penentuan kelayakan bantuan sembako menggunakan algoritma K-Means. Data yang digunakan untuk menentukan kelayakan yaitu data yang diambil dari data masyarakat di Desa Bangun Purba.

2. Metode K-Means memanfaatkan pengelompokan data untuk menghasilkan probabilitas setiap kriteria untuk class yang berbeda, sehingga nilai-nilai probabilitas dari kriteria tersebut dapat dioptimalkan untuk menentukan kelayakan penerima bantuan bedah rumah berdasarkan proses klasifikasi yang dilakukan oleh metode K-Means itu sendiri.
3. Aplikasi data mining kelayakan bantuan bedah rumah ini dibangun menggunakan bahasa pemrograman PHP dan database MySQL.

## Saran

Berdasarkan uraian dan pembahasan pada bab-bab sebelumnya, maka dapat ditarik saran-saran sebagai berikut:

1. Dengan adanya penelitian ini dapat menjadi acuan kepala desa untuk mengambil kebijakan tahap selanjutnya sesuai dengan pengetahuan yang dihasilkan oleh data mining metode K-Means ini agar penilaian prioritas penduduk penerima bantuan bedah rumah bisa lebih berkualitas dan efektif.
2. Penelitian ini dapat dijadikan sebagai salah satu referensi bagi aparat desa untuk digunakan sebagai acuan dalam pengambilan keputusan untuk menentukan kelayakan penerima bantuan bedah rumah

## VI. Daftar Pustaka

Chamida, M. A., Susanto, A., & Latubessy, A. (2021). Analisa User

- Acceptance Testing Terhadap Sistem Informasi Pengelolaan Bedah Rumah Di Dinas Perumahan Rakyat Dan Kawasan Permukiman Kabupaten Jepara. *Indonesian Journal of Technology, Informatics and Science (IJTIS)*, 3(1), 36–41. 1
- Damanik, Y. F. S., Yani, R., Sumarno, S., Gunawan, I., Hartama, D., & Kirana, I. O. (2021). Penerapan Data Mining Untuk Pengelompokan Penyebaran Covid-19 Di Sumatera Utara Menggunakan Algoritma K-Means. *Jurnal Ilmu Komputer Dan Informatika*, 1(2), 109–132.
- Dewi, S. P., Nurwati, N., & Rahayu, E. (2022). Penerapan Data Mining Untuk Prediksi Penjualan Produk Terlaris Menggunakan Metode K-Nearest Neighbor. *Building of Informatics, Technology and Science (BITS)*, 3(4), 639–648.
- Fetriany, I., & Sobari, I. A. (2020). jurnal perancangan apk\_UN. *Jurnal Komputa, Informatika*, 9(2).
- Kusnadi, Y., & Putri, M. S. (2021). Clustering Menggunakan Metode K-Means Untuk Menentukan Prioritas Penerima Bantuan Bedah Rumah (Studi Kasus : Desa Ciomas Bogor). *Jurnal Teknologi Informatika Dan Komputer*, 7(1), 17–24.
- Nengsih, Y. G. (2020). Sistem Informasi Penjualan Jilbab Pada Toko Karunia Bukittinggi Dengan Bahasa Pemrograman Php Dan Mysql. *Jurnal Pendidikan Teknologi Informasi (JUKANTI)*, 3(1), 21–28.
- Noviantoro, A., Silviana, A. B., Fitriani, R. R., & Permatasari, H. P. (2022). Rancangan Dan Implementasi Aplikasi Sewa Lapangan Badminton Wilayah Depok Berbasis Web Agung. *Jurnal Teknik Dan Science*, 1(2), 88–103.
- Nur, A., & Daeli, Y. (2024). Pengembangan Rekayasa Perangkat Lunak Online Sistem E-Recruitment pegawai PT.FM Global Logistics Berbasis web. *Journal Of Informatics And Busisnes*, 02(03), 468–479.
- Pangestu, A., & Ridwan, T. (2022). Penerapan Data Mining Menggunakan Algoritma K-Means Pengelompokan Pelanggan Berdasarkan Kubikasi Air Terjual Menggunakan Weka. *JUST IT : Jurnal Sistem Informasi, Teknologi Informasi Dan Komputer*, 11(3), 67–71.
- Pasaribu, F. M. (2022). Penerapan Metode Promethee II Dalam Sistem Pendukung Keputusan Pemilihan Penerima Bantuan Bedah Rumah Di Desa Hutapaung. *Konferensi Nasional Teknologi Informasi Dan Komputer*, 6(1), 436–447.
- Prastiwi, H., Jeny Pricilia, & Errissy Rasywir. (2022). Implementasi Data Mining Untuk Menentuksn Persediaan Stok Barang Di Mini Market Menggunakan Metode K-Means Clustering. *Jurnal Informatika Dan Rekayasa Komputer(JAKAKOM)*, 2(1), 141–148.

- Premana, A., Wijaya, A. P., Yono, R. R., & Hayati, S. N. (2022). Media Pembelajaran Pengenalan Bahasa Pemrograman Pada Anak Usia Dini Berbasis Game. *Tekinfor: Jurnal Bidang Teknik Industri Dan Teknik Informatika*, 23(2), 66–75.
- Rahayu, D. S., Afifah, J., & Intan, S. (2023). Classification of Diabetes Mellitus Using C4 . 5 Algorithm , Support Vector Machine ( SVM ) and Linear Regression Klasifikasi Penyakit Diabetes Melitus Menggunakan Algoritma C4 . 5 , Support Vector Machine ( SVM ) dan Regresi Linear. *SENTIMAS: Seminar Nasional Penelitian Dan Pengabdian Masyarakat*, 1(1 SE-), 56–63.
- Rudjiono, D., & Saputro, H. (2021). Pengembangan Desain Website Sebagai Media Informas Dan Promosi (Studi Kasus: PT.Nada Surya Tunggal Kecamatan Pringapus). *Pixel :Jurnal Ilmiah Komputer Grafis*, 13(2), 56–66.
- Silalahi, N. (2020). Penentuan Strategi Promosi Universitas Budi Darma Menggunakan Algoritma K-Means Clustering. *TIN: Terapan Informatika Nusantara*, 1(1), 40–46.
- Srirahayu, A., & Pribadie, L. S. (2023). Review Paper Data Mining Klasifikasi Data Mining. *Jurnal Ilmiah Informatika Global*, 14(1).
- Suhartini, S., Sadali, M., & Putra, Y. K. (2020). Sistem Informasi Berbasis Web Sma Al- Mukhtariyah Mamben Lauk Berbasis Php Dan Mysql. *Jurnal Informatika Dan Teknologi*, 3(1), 79–83.
- Yulianingsih, E., Panjaitan, F., Fatmasari, Yudiastuti, H., & Oktarini, N. A. (2022). Introduction and Training of Database Systems for Vocational High School Students in Palembang City. *Jurnal Pengabdian Pancasila (JPP)*, 1(2), 69–76.